



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/821,908	03/30/2001	Jonathan Stern	2937.1000-004	7680
21005	7590	08/16/2004	EXAMINER	
HAMILTON, BROOK, SMITH & REYNOLDS, P.C. 530 VIRGINIA ROAD P.O. BOX 9133 CONCORD, MA 01742-9133			HARRELL, ROBERT B	
			ART UNIT	PAPER NUMBER
			2142	

DATE MAILED: 08/16/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No.

09/821,908

Applicant(s)

STERN ET AL.

Examiner

Robert B. Harrell

Art Unit

2142

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 30 March 2001.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-25 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-25 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 30 March 2001 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date since 3/30/2001.
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☒ Other: see attached Office Action.

Art Unit: 2142

1. Claims 1-25 are presented for examination.
2. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.
3. The form and legal phraseology often used in patent claims, such as "means" should be avoided in the Abstract. (see MPEP 608.01(b)).
4. Page 5 (line 28) recited "5,6, and 7"; examiner is unable to associate these with the figures nor is able to associate the "application" to which this line makes reference. Thus, correction is required; and, also, the applicant should also use this period for response to thoroughly and very closely proof read and review the whole of the application for correct correlation between reference numerals in the textual portion of the Specification and Drawings along with any minor spelling errors, general typographical errors, accuracy, and clarity of meaning in the Specification, Drawings, and specifically claims. Minor typographical errors could render a Patent unenforceable and so the applicant is strongly encouraged to aid in this endeavor.
5. Use of active hyperlink and/or other forms of browser executable code is improper (see MPEP 608.01) and must be removed (see page 3 (lines 8-11)) as one example, all others must also be removed. However, lines 16-19 of page 3 is acceptable (no "http://"). The reason being many OCR softwares will automatically associate (convert) such as an active (clickable) URL. Also, due to the dynamic nature of such links, they tend to become disabled rather than not in a relatively short time compared to the life of a Patent. It is suggested to use the following:  
  
"<protocol : <domain> <directory> <file>  
where protocol may be of the type http, domain is a domain name of a site and directory the name of the directory in which a file so named is located."
6. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this action:  
  
**A person shall be entitled to a patent unless -**  
  
**(b) the invention was patented or described in a printed publication in this or foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.**
7. Claims 1-25 are rejected under 35 U.S.C. 102 (b) as being anticipated by Sahami et al. "SONIA: A Service for Organizing Networked Information Autonomously" (pages 200-209) as provided by the applicant in January 16, 2003 by the applicant.
8. Prior to addressing the grounds of the rejections below, should this application ever be the subject of public review by third parties not so versed with the technology (i.e., access to IFW through Public PAIR (as found on <http://portal.uspto.gov/external/portal/pair>)), the following additional indicia in this examiner's Office Action is an aid to refer attention to

Art Unit: 2142

relevant and helpful elements, figures, and/or text upon which the examiner relies to support his position. Thus, the following citations are neither all-inclusive nor all-exclusive in nature as the whole of the reference is cited and relied upon in this action.

9. Per claim 1, Sahami taught a method (e.g., see page 203 (figure 2) and page 205 (right second line after "Controlled Experiments" "methods")) for searching (e.g., see page 200 ("Abstract" (line 3 "search"))) for people and organization information (e.g., see page 206 (left ("Usage Scenario One" ("Hector Garcia-Molina") is a person and the organization is "the Stanford Digital Libraries Project"))); also see page 207 (Table 4 (far left (third box from the top (first line "university"))))) on Web pages (e.g., see page 206 (left (fourth to the last line "Web pages"))) in a global computer network (see figure 1 on page 201 as described on that page on the right in the last paragraph; also note "AltaVista" in figure 1, "Yahoo" on page 200 (right (second paragraph (first line))) and "World Wide Web" on the second line of the Abstract on page 200, all anticipating the "Internet" which has been (since mid 1960s) a World Wide, hence "global", computer network) comprising the steps of:

a) accessing a Web site (e.g., see page 202 (left ("Document retrieval and parsing" [note the apostrophe in "(e.g., URL's" (seventh to the last line of page 202 (left))) is located between the "L" and "s" suggestive of a Web site] of potential interest (e.g., see page 206 (right (last paragraph (line 7 "interested" and thus a potential input for the SONIA method))))), the Web site having a plurality of Web pages (called "documents") (see figure 2 on page 2-3 (far left after "Parallelized Crawler" showing five pages of text within a dotted area titled "Document Text"));

b) determining (e.g., see figure 2 (bottom right "Clusterer") and page 204 starting on the left "Clustering" to the last line of the right side of that page) a subset (e.g., see page 203 (figure 2 (far bottom right "Subset"))) of the plurality of Web pages (obtained from the far left of figure 2) to process (e.g., see figure 2 (far lower right "Descriptor Extractor" as covered on page 205 (top left)); and

c) for each Web page in the subset, (i) determining types of contents found on the Web page (e.g., descriptors and/or topical descriptors for terms in each subset as covered on page 205 (starting with top left) and on page 207 (Table 4 "Feasible Topics"), and (ii) based on the determined content types (i.e., University of Stafford per page 206 ("Usage Scenario One") as found in table 4 of page 207), enabling extraction of people and organization information from the Web page (e.g., see page 206 (left ("Usage Scenario One" ("Hector Garcia-Molina") is a person and the organization is "the Stanford Digital Libraries Project"))); also see page 207 (Table 4 (far left (third box from the top (first line "university"))))).

10. Per claim 2, see page 203 (right (last line "Bayesian networks")) where "content owner" would be the University of Stafford that held the Web Page as listed on page 207 in Table 4 (Top right first box).

11. Per claim 3, 4, 5, 8 and 9, see page 206 (left (fifth to the last line from the bottom "crawler")). Web Crawlers, (aka Spiders page 202 (left (fourth to the last line))), "crawl" (functionally it is a scan process, there is no physical activity in the sense of something moving about) the World Wide Web (see page 200 (Abstract (line 2))) for data

Art Unit: 2142

and subsequent links to more Web Sites to which Spider my "crawl" and obtain more Web Pages; with yet more URLs. Since the applicant is well versed with "crawlers" (see page 6 (line 4) of this application) examiner takes *Official Notice* of web crawler functions.

12. Per claims 6, 7 and 10, Sahami gave an example on page 207 (left ("Usage Scenario Two")). Such would have resulted in a news article of a press release directed to a question about life on one of Saturn's moon Titan. Furthermore, document content is "printed subject matter" and also a selective design choice as to content based on the user's interest such as John Hancock's "signature" on the Declaration of Independence (i.e., the student is searching parts of the United States' history).

13. Per claims 11 and 12, see page 206 (left (fourth and third to the last line)).

14. Per claim 13, such are anticipated crawler/spider variables and anticipated parameters to web crawlers.

15. Per claims 14-25, these claims do not teach or defined above the correspondingly rejected claims given above, and are thus rejected for the same reasons given above since the above citations into Sahami are functionally equivalent to the means provided in the specification of this application and as covered by the scope of the defined invention. However, as for "database", of claim 14, see page 202 (left (line 21)); also, Internet sites such as AltaVista (figure 2) contained a database of URLs.

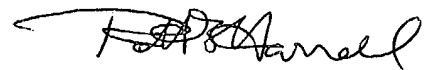
16. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Najork et al. (US 6418433 B1) taught a Web Crawler system in which, among many other things, was shown in figure 3 (164) the gathering of URLs from snagged Web Pages for addition to the list of URLs to visit.

17. A shortened statutory period for response to this action is set to expire **3 (three) months** and 0 (zero) days from the data of this letter. Failure to respond within the period for response will cause the application to become abandoned (see MPEP 710.02, 710.02(b)).

18. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Robert B. Harrell whose telephone number is (703) 305-9692. The examiner can normally be reached Monday thru Friday from 5:30 am to 2:00 pm and on weekends from 6:00 am to 12 noon Eastern Standard Time.

19. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jack B. Harvey, can be reached on (703) 308-9705. The fax phone numbers for the Group are (703) 746-7238 for After-Final, (703) 746-7239 for Official Papers, and (703) 746-7240 for Non-Official and Draft papers.

20. Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group receptionist whose telephone number is (703) 305-9600.



**ROBERT B. HARRELL**  
**PRIMARY EXAMINER**  
Art Unit 2142